On-Line Character Recognition Adaptively Controlled by Handwriting Quality

Masahiko HAMANAKA and Keiji YAMADA
Computer & Communication Media Research, NEC Corporation
4-1-1 Miyazaki, Miyamae-ku, Kawasaki, Kanagawa 216-8555, Japan
E-mail: {m-hamanaka@az, kg-yamada@cp}.jp.nec.com

On-line character recognition which can adapt to handwriting quality is proposed. In character recognition, it is difficult to recognize both clearly and roughly written characters accurately. For Japanese characters, the number of strokes is often slightly varied when characters are written roughly. In a previous method, the ranges of the number of strokes were set widely enough for recognition; however, these ranges were not optimal for clearly written characters. The proposed method controls a distribution model of the number of strokes adaptively according to handwriting quality, and it uses this model for pre-candidate selection and fine classification. Recognition experiments demonstrated that the proposed method has greater recognition accuracy and speed than the previous method. In particular, accuracy was improved from 91.4% to 94.3% and speed was increased by about 50% when recognizing clearly written data.

1 Introduction

On-line handwriting recognition interests mobile-computer users as an easy input method. Easy input requires highly accurate and speedy recognition, while permitting many kinds of variations of handwritten characters: for example, the position and size of characters, the shape of characters, and the number and order of strokes used to write them. Some people write characters carefully and clearly, while other people write them rapidly and roughly with distorted shapes and the incorrect number and order of strokes. On-line character recognition therefore needs to be able to recognize both clearly and roughly written characters accurately.

Many methods have been developed for on-line character recognition [1,2], and some of them have been improved to the point of being able to recognize roughly written characters. However, these improvements have decreased their accuracy in recognizing clearly written characters. The major problem is therefore how to increase recognition accuracy for both clearly written characters and roughly written ones. This is a serious problem for Japanese character recognition in particular because of the complexity of the characters used in Japanese.

Writers of the Japanese language employ a large set of Kanji, and two syllabaries, Hiragana and Katakana, as well as alphanumerics and other common

In: L.R.B. Schomaker and L.G. Vuurpijl (Eds.), Proceedings of the Seventh International Workshop on Frontiers in Handwriting Recognition, September 11-13 2000, Amsterdam, ISBN 90-76942-01-3, Nijmegen: International Unipen Foundation, pp 23-32

daily symbols. Kanji characters are Chinese ideographs; 2,965 are designated in the first level of the Japanese Industry Standard (JIS) code set. Hiragana and Katakana both include approximately 80 cursive symbols for consonant and vowel combinations. In particular, Kanji characters are complex symbols with many variations in the number and order of strokes.

The authors earlier proposed on-line Japanese character recognition based on a pattern-matching method used in OCRs [3]. A simple pattern-matching method sometimes misrecognizes characters because it does not use the number of strokes, though it can recognize characters irrespective of both the number and order of strokes. To solve this problem, the authors introduced a pre-candidate selection using the number and length of strokes [4]. This selection improved both its recognition accuracy and speed.

Although the number of strokes rarely changes when characters are written clearly, it changes variously when characters are written roughly. The permission ranges of the number of strokes for pre-candidate selection are therefore set widely enough so that characters written roughly can be recognized. Conversely, this condition is not optimal for characters written clearly. In this paper, the authors focus on the number of strokes because it differs widely from character to character and can be controlled easily. The authors propose to control a distribution model of the number of strokes according to handwriting quality. In addition, they introduce the distribution model of the number of strokes not only into the pre-candidate selection but also into the fine classification.

The rest of the paper is organized as follows: distributions of the number of strokes are analyzed and distribution models dependent on handwriting quality are proposed in Section 2. How to select the handwriting quality is discussed in Section 3. Section 4 describes the recognition algorithm and Section 5 shows experimental results.

2 Distributions of the Number of Strokes

2.1 Pre-candidate selection using the number of strokes

For pre-candidate selection by using the number of strokes, each category k possesses a lower bound, $N^{low}(k)$, and an upper bound, $N^{up}(k)$, of the permission range of the number of strokes. If the number of strokes, N_{in}, extracted from an unknown input satisfies the condition

$$N^{low}(k) \leq N_{in} \leq N^{up}(k), \tag{1}$$

then the category k remains as a candidate to be dealt with in the next classification stage (Fig. 1).

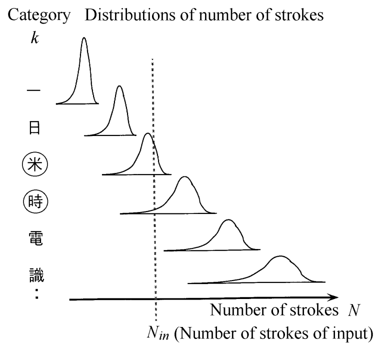

Figure 1: Pre-candidate selection using the number of strokes.

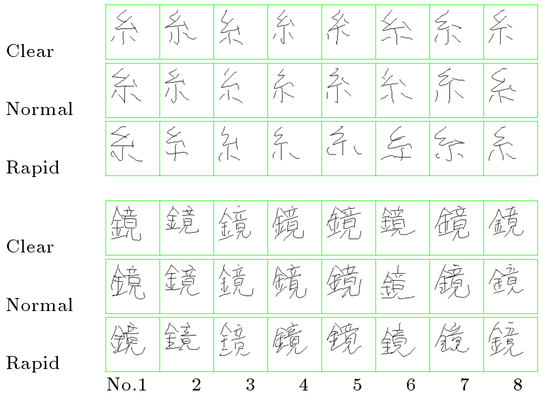

Figure 2: Examples of collected data.

Previously, the lower bound, $N^{low}(k)$, and upper bound, $N^{up}(k)$, of the number of strokes were determined from the minimum, $N^{min}(k)$, and maximum, $N^{max}(k)$, of the number of strokes by using all learning data as $N^{low}(k) = N^{min}(k) - a$ and $N^{up}(k) = N^{max}(k) + a$, where a is a parameter for adjusting the permission range (usually $a = 1$). Here, the narrower the permission range is, the faster the recognition speed is. And if categories that could be mistakenly recognized instead of correct categories are excluded by pre-candidate selection, the recognition rate will be higher. In practice, the permission range for pre-candidate selection needs to be made as narrow as possible, so as to exclude incorrect categories, while the correct categories are retained.

2.2 Analysis of distributions of the number of strokes

As an on-line database for analyzing distributions of the number of strokes, the authors collected a total of 24 samples by eight writers; each writer wrote three samples under conditions of handwriting quality: “clear”, “normal”, and “rapid”. Each sample includes 342 characters: 71 Hiragana, 71 Katakana, 100 simple Kanji, and 100 complex Kanji characters. Figure 2 shows examples of the collected data. Most data written clearly are written with the correct number of strokes, and most data written roughly tend to be written with connected strokes.

Figure 3 shows distributions of the number of strokes of categories which possess the same standard number of strokes; the distributions are calculated corresponding to each handwriting quality. The standard number of strokes is determined as the number of strokes which gives the maximum frequency by

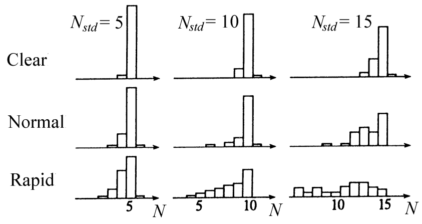

Figure 3: Distributions of the number of strokes for each standard number of strokes.

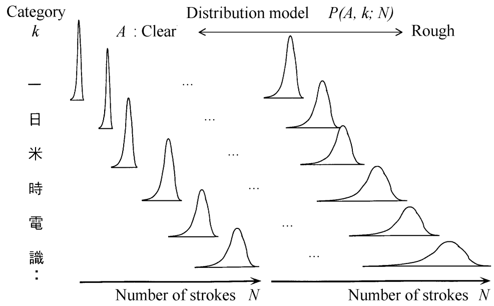

Figure 4: Distribution model of the number of strokes for categories.

using distributions of characters written clearly. The figure suggests the larger the standard number of strokes is, the wider the distribution is. And the lower the handwriting quality is, the wider the distribution is. In addition, for characters written roughly, the number of strokes which gives the maximum frequency is sometimes smaller than the standard number of strokes. It is therefore expected that recognition will be effective if the distribution model of the number of strokes is changed adaptively according to handwriting quality.

2.3 Distribution models of the number of strokes

Actually, the distribution of the number of strokes depends on the categories. When a distribution model of the number of strokes for each category is determined according to handwriting quality, it can be expressed as $P(A, k; N)$, where A is the handwriting quality, k is the category, and N is the number of strokes (Fig. 4). This model is considered to be the best for recognition. However, a large number of data is necessary in order to determine the distribution model for each category and each handwriting quality, and the memory size of a dictionary is not small. For example, when there are 4000 categories and three levels of handwriting quality, and if 100 data are used for each distribution, then a total of 1.2 million data are necessary. And if 10 bytes are used for each distribution, a total of 120K bytes of memory are necessary.

When a distribution model of the number of strokes for each standard number of strokes is determined, it can be expressed as $P'(A, N_{std}(k); N)$, where each category possesses only its standard number of strokes $N_{std}(k)$ (Fig. 5). Although it is assumed that categories which possess the same standard number of strokes have the same distribution of the number of strokes, fewer learning data and less dictionary memory are necessary. For example, when

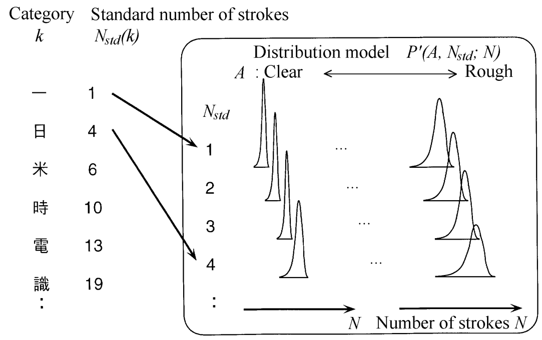

Figure 5: Distribution model of the number of strokes for standard numbers of strokes.

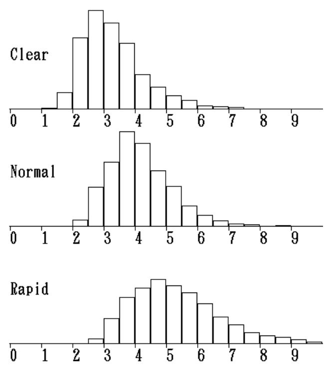

Figure 6: Distributions of the average distance between sampling points.

up to 30 standard numbers of strokes are dealt with, only 9000 learning data and about 5K bytes of memory are necessary. In this paper, the distribution model $P'(A, N_{std}(k); N)$ according to the standard number of strokes is used.

3 Selection of Handwriting Quality

Handwriting quality can be selected by manual setting or automatic detection. For manual setting, before recognition, writers select the handwriting quality with which they will write characters. Automatic detection of handwriting quality from unknown input data is difficult; however, one detection method is to use writing speeds. The writing speed can be calculated from the average distance between sampling points.

Figure 6 shows distributions of the average distance between sampling points using the data collected under three conditions of handwriting quality. When characters are written rapidly, the writing speed tends to be fast, so the average distance between sampling points is large. In practice, not all data were written according to the condition of handwriting quality. Although there are overlaps among distributions, handwriting quality can be detected from the average distance between sampling points. Regarding this data, for example, if handwriting quality is decided as “clear” when the average distance between sampling points is under 3.0, “normal” when from 3.0 to 4.5, and “rough” when over 4.5, then 44% of clearly written data, 57% of normally written data, and 69% of rapidly written data are determined correctly respectively.

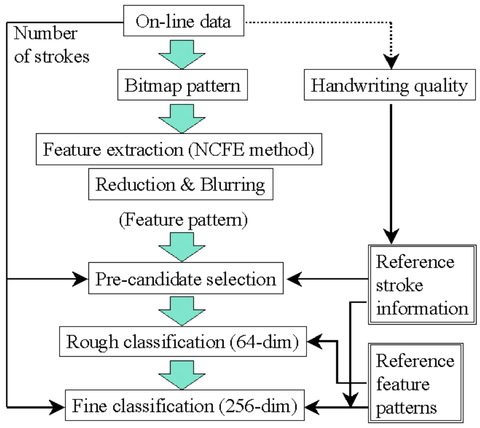

Figure 7: On-line recognition system.

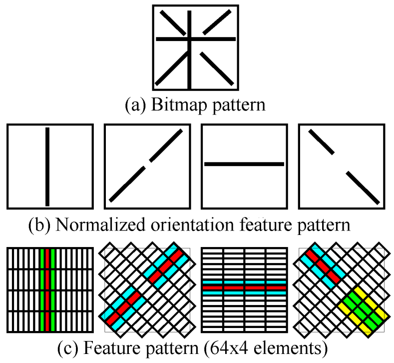

Figure 8: An example of feature extraction.

4 Recognition Algorithm

Figure 7 outlines on-line character recognition based on the flexible-pattern-matching (FPM) technique[3] with stroke information adaptively controlled by handwriting quality. On-line data of a character are first normalized according to size, resampled and smoothed, and then transformed to a bitmap pattern (Fig. 8(a)). This bitmap pattern is transformed into a normalized orientation feature pattern (Fig. 8(b)) by using normalization-cooperative feature extraction (NCFE)[5], in which an orientation feature pattern is extracted from the bitmap pattern and then transformed nonlinearly by using line-density normalization functions. It is then resolved to a 256-element feature pattern and blurred (Fig. 8(c)). Here the feature pattern elements are summed as a complexity feature, which is approximately in proportion to the total length of strokes. And the 256-element feature pattern is normalized such that the summation of the normalized feature pattern elements is a constant. In addition, a 64-element feature pattern is generated from the 256-element feature pattern by averaging the four neighboring elements.

As a dictionary, the reference stroke information includes, first, the lower bound, $N^{low}(A, N_{std}(k))$, the upper bound, $N^{up}(A, N_{std}(k))$, and the distribution model $P'(A, N_{std}(k); N)$ of the number of strokes for each handwriting quality A and each standard number of strokes, $N_{std}(k)$; secondly, it includes the standard number of strokes, $N_{std}(k)$, for each category k; thirdly, it includes the lower bound $C^{low}(k)$ and the upper bound $C^{up}(k)$ of the complexity feature (the total length of strokes) for each category k. The reference feature patterns consist of 64-element reference feature patterns and 256-element reference feature patterns.

The classification process has three stages: pre-candidate selection using the number of strokes and the complexity feature, rough classification using the 64-element feature pattern, and fine classification using the 256-element feature pattern and the number of strokes. During the pre-candidate selection, if the number of strokes N_{in} extracted from an unknown input satisfies the condition

$$N^{low}(A, N_{std}(k)) \leq N_{in} \leq N^{up}(A, N_{std}(k)), \tag{2}$$

and if the complexity feature C_{in} extracted from an unknown input satisfies the condition $C^{low}(k) \leq C_{in} \leq C^{up}(k)$, then the category k remains as a candidate.

During the rough classification, the 64-element feature pattern is matched with the 64-element reference feature patterns by linear pattern matching using a city-block distance. This process reduces the number of candidates to 10. During the fine classification, the 256-element feature pattern is matched with the 256-element reference feature patterns of the candidates by nonlinear pattern matching based on dynamic programming. In the 256-element feature pattern, the distance between the input and category k is expressed as $D_f(k)$, then a modified distance $D_m(k)$ taking account of the number of strokes is calculated from

$$D_m(k) = D_f(k) - w\ P'(A, N_{std}(k); N_{in}), \tag{3}$$

where w is a weight parameter.

5 Experiments

5.1 Learning data and dictionary

As an on-line database for learning, the authors collected a total of 51 samples, freely written by 51 untrained writers. Each sample includes the 2,965 Kanji characters in the first level of the JIS code set and 232 non-Kanji characters (71 Hiragana characters, 71 Katakana characters, 62 alphanumerics, and 28 other symbols). After 245 miswritten data were removed, a total of 162,802 remained.

Learning data for each category k were used to determine the standard number of strokes $N_{std}(k)$, the lower bound $C^{low}(k)$ and the upper bound $C^{up}(k)$ of the complexity feature, and the reference feature patterns automatically. However, since the learning data were written freely, the other features dependent on handwriting quality in the reference stroke information can not

Table 1: Settings for the ranges of the number of strokes.
(Upper line: upper bound $N^{up}(A, N_{std}(k))$; Lower line: lower bound $N^{low}(A, N_{std}(k))$)

Handwriting quality A	Standard number of strokes $N_{std}(k)$															
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	...
Clear	1	2	3	4	5	7	8	9	10	11	12	13	14	15	16	...
	1	2	3	3	4	5	6	7	8	9	10	10	11	12	13	...
Normal	1	3	4	5	6	7	8	9	10	11	12	13	14	15	16	...
	1	1	2	3	3	4	5	5	6	7	8	9	9	10	10	...
Rough	1	3	4	5	6	7	8	9	10	11	12	12	13	14	15	...
	1	1	1	2	2	3	3	4	4	5	5	6	6	7	7	...

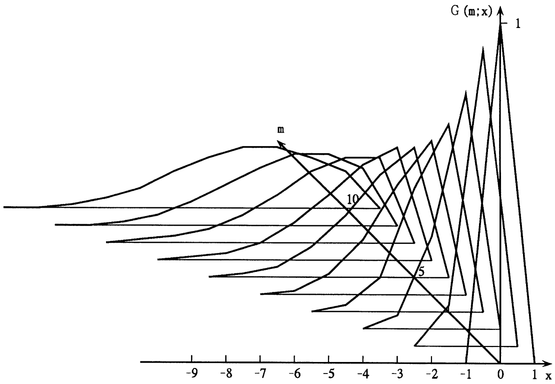

Figure 9: Distribution function $G(m; x)$.

be determined from the learning data. They were therefore determined empirically in the following way. First, the lower bound, $N^{low}(A, N_{std}(k))$, and upper bound, $N^{up}(A, N_{std}(k))$, of the number of strokes according to handwriting quality and the standard number of strokes were determined as listed in Table 1. Next, the distribution model was determined as

$$P'(A, N_{std}(k); N) \;=\; G(N_{std}(k) - N^{low}(A, N_{std}(k)); N - N_{std}(k)), \qquad (4)$$

where $G(m; x)$ is a set of distribution functions, as shown in Fig. 9, which were determined by using normal distributions. When using Eqn. (4), the distribution model $P'(A, N_{std}(k); N)$ becomes the same for categories k, if the differences between the standard number of strokes, $N_{std}(k)$, and the lower bound, $N^{low}(A, N_{std}(k))$, are the same.

5.2 Experimental results

As an on-line database for testing, the on-line database collected under conditions of handwriting quality mentioned in Section 2.2 was used. Three recognition methods were compared to evaluate their respective effectiveness: 'Exp-1' is the previous recognition method with pre-candidate selection using Eqn. (1) and fine classification using distance $D_f(k)$; 'Exp-2' is a proposed method with pre-candidate selection using Eqn. (2) and fine classification using distance $D_f(k)$; and 'Exp-3' is another proposed method with pre-candidate selection using Eqn. (2) and fine classification using distance $D_m(k)$.

Table 2 lists recognition rates and cumulative rates of the top 10 candidates in 'Exp-1' and 'Exp-2'. The recognition rates are counted when the first candidate category is correct or is one of the same shape characters of the correct category. Each data set was tested according to conditions of handwriting quality A: "clear", "normal", and "rough". When handwriting quality A perfectly coincides with the actual handwriting quality of data, recognition rates increased; in particular, the recognition rate of clearly written data improved from 91.37% to 93.24%. And when the handwriting quality A is lower than the actual handwriting quality of data, recognition rates were better than those of 'Exp-1'. On the other hand, when it is higher, recognition rates dropped considerably compared with those of 'Exp-1'. Since correct categories tend to be excluded from candidates by pre-candidate selection, cumulative recognition rates also dropped. The handwriting quality A therefore needs to be made equal to or at least lower than the actual handwriting quality of data.

Table 3 lists recognition rates measured in 'Exp-3'. Each data set was tested with the most suitable condition of handwriting quality. When weight parameter w was adjusted, recognition rates increased by about 1 - 2%. For example, for clearly written data, the recognition rate increased from 93.24% to 94.33% when w was 900.

Recognition speed was also investigated. When the recognition time measured in 'Exp-1' was 100, the times measured in 'Exp-3' for recognizing data of handwriting quality "clear", "normal" and "rough" were about 68, 85 and 103, respectively. When data of handwriting quality "clear" were recognized, recognition speed increased to about 1.5 times that of 'Exp-1'.

6 Conclusion

The authors have proposed on-line character recognition adaptively controlled by handwriting quality for accurately recognizing both clearly and roughly written characters. The method controls a distribution model of the number of strokes according to handwriting quality. It then uses the model for pre-

Table 2: Recognition rates (%) from 'Exp-1' and 'Exp-2'.
Numbers in parentheses show cumulative rates within the top 10.

Data	'Exp-1'	'Exp-2': Handwriting quality A Clear	Normal	Rough
Clear	91.37 (98.43)	93.24 (98.50)	92.00 (98.61)	91.41 (97.88)
Normal	88.16 (98.06)	86.22 (93.31)	88.41 (97.95)	88.23 (97.77)
Rapid	74.27 (91.85)	62.79 (73.68)	72.19 (88.41)	74.78 (92.51)

Table 3: Recognition rates (%) from 'Exp-3'.

Data	Handwriting quality A	Weight w 0	300	500	700	900	1100
Clear	Clear	93.24	94.12	94.26	94.30	94.33	94.19
Normal	Normal	88.41	89.77	90.10	90.28	90.46	90.53
Rapid	Rough	74.78	76.17	76.39	76.61	76.61	76.57

candidate selection and fine classification. Although a distribution model of the standard number of strokes was used approximately, experiments demonstrated the proposed method improved both accuracy and speed of recognition. In particular, for clearly written data, the recognition rate increased from 91.4% to 94.3% and the recognition speed increased by about 50%.

Further research will be focussed on, for example, adjusting distribution models of the number of strokes by using learning data, promoting adaptive learning of distribution models, and developing automatic detection of handwriting quality.

References

1. C. C. Tappert, C. Y. Suen, and T. Wakahara, "The State of the Art in On-Line Handwriting Recognition," *IEEE Trans. PAMI*, Vol.12, No.8, pp.787-808 (1990).
2. M. Nakagawa, "Non-keyboard Input of Japanese Text On-line Recognition of Handwritten Characters as the Most Hopeful Approach," *J. Information Processing*, Vol.13, No.1, pp.15-34 (1990).
3. M. Hamanaka, K. Yamada, and J. Tsukumo, "On-Line Japanese Character Recognition Experiments By an Off-line Method Based on Normalization-Cooperative Feature Extraction," *Proc. 2nd ICDAR*, pp.204-207 (1993).
4. M. Hamanaka and K. Yamada, "On-Line Japanese Character Recognition Based on Flexible Pattern Matching with Stroke Information," *Proc. 5th IWFHR*, pp.317-320 (1996).
5. M. Hamanaka, K. Yamada, and J. Tsukumo, "Handprinted Kanji Character Recognition Using Normalization-Cooperated Feature Extraction," *Proc. 3rd IWFHR*, pp.343-348 (1993).